DOCUMENT RESUME

ED 216 014 TM 820 151

AUTHOR Zyskowski, Gloria
TITLE How Not to Conduct a Naturalistic Evaluation: Rueful Reminiscences of a Museum Evaluator.
PUB DATE Oct 81
NOTE 12p.; Paper presented at the Annual Meeting of the Evaluation Network (Austin, TX, October 1-3, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Methods; *Institutional Evaluation; *Measurement Techniques; *Museums
IDENTIFIERS Evaluation Problems; Evaluation Utilization; *Naturalistic Evaluation

ABSTRACT

While evaluating the effectiveness of a portion of the public information system at the Toledo Art Museum, nonclassical approaches to data-gathering presented problems in measuring program aspects which are not easily quantified. There is a necessity to clarify what kinds of information can validly be gathered with naturalistic techniques for an audience unfamiliar with evaluation processes but having specific evaluation needs within their organizational activities. Curators and museum educators are seen as looking for evidence of the beneficial effects of the public information program while sound evaluation techniques and professional standards must be maintained. The difficulties of nonclassical evaluation are discussed in terms of problem formulation, the iterative framework for refinement of questions and the participation of the evaluation audience. Precautions for use of naturalistic evaluation are outlined and within the limitations of the techniques in non-traditional settings a payoff is suggested; the kinds of data gathered and reports submitted are often of immediate and practical use to the evaluation audience. Nontraditional forms of evaluation are seen as likely to become increasingly pervasive. (Author/CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED216014

How Not to Conduct a Naturalistic Evaluation: Rueful Reminiscences of a Museum Evaluator

BY

Gloria Zyskowski, Ph.D.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
G. Zyskowski
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A paper presented at the
Evaluation Network Conference in Austin, Texas
October 1-3, 1981

TM 820 151

WHILE CONDUCTING AN EVALUATION OF A PORTION OF THE PUBLIC INFORMATION SYSTEM AT THE TOLEDO MUSEUM OF ART AND EMPLOYING A NATURALISTIC APPROACH TO DETERMINE ITS EFFECTIVENESS, THIS RESEARCHER BECAME AWARE OF THE PROBLEMS INHERENT IN MEASURING THINGS WHICH ARE NOT EASILY QUANTIFIABLE. IN ADDITION, THIS RESEARCHER GAINED AN APPRECIATION FOR THE NECESSITY OF CLARIFYING WHAT KINDS OF INFORMATION CAN VALIDLY BE GATHERED FROM SUCH TECHNIQUES TO AN AUDIENCE (CURATORS AND MUSEUM EDUCATORS) RELATIVELY UNSOPHISTICATED IN EVALUATION PROCESSES BUT NEVERTHELESS EXPECTING TO FIND ANSWERS TO ALL OF THEIR QUESTIONS.

THIS PAPER CONTAINS A DISCUSSION OF THE PROBLEMS ENCOUNTERED WHEN NON-CLASSICAL APPROACHES TO DATA-GATHERING ARE EMPLOYED, WITH ILLUSTRATION FROM A MUSEUM EXPERIENCE. IN ADDITION, GENERAL ADVICE IS OFFERED TO RESEARCHERS INTERESTED IN USING NATURALISTIC/RESPONSIVE TECHNIQUES IN ADDRESSING PARTICULAR EVALUATION QUESTIONS.

EXPOSITION OF THE ISSUE

IF WE WERE IN LOGIC 101 DISCUSSING VARIOUS TECHNIQUES ASSOCIATED WITH THE SCIENCE OF ARGUMENTATION, THE PHENOMENON WOULD BE TERMED "THE BANDWAGON EFFECT". THAT IS, "EVERYBODY'S DOING IT, SO I BETTER DO IT TOO IF I DON'T WANT TO BE LEFT BEHIND." JUST WHAT IS IT THAT EVERYBODY SEEMS TO BE DOING? AMONG THE LESS PROSAIC ANSWERS TO THAT QUESTION, AND THE ONE ADDRESSED IN THIS PAPER, IS "EVALUATION".

TODAY, PROGRAMS OTHER THAN THOSE ORIGINATING FROM THE FIELDS OF GOVERNMENT, EDUCATION, AND THE NATURAL SCIENCES ARE CANDIDATES FOR AN EVALUATOR'S CONSIDERATION. VARIOUS ATYPICAL ORGANIZATIONS ARE JUMPING

ON THE FORMATIVE AND SUMMATIVE BANDWAGONS, BRINGING WITH THEM SETS OF UNIQUE PROBLEMS. FOR EXAMPLE, ART MUSEUMS PRESENTLY ARE SERVING AS THE FOCAL POINTS FOR MANY RESEARCH ACTIVITIES. CURATORS AND MUSEUM EDUCATORS WHO HAVE DESIGNED PUBLIC INFORMATION PROGRAMS ARE LOOKING FOR EVIDENCE THAT WHAT THEY ARE DOING HAS SOME BENEFICIAL EFFECTS, EITHER SHORT-TERM OR LONG-TERM, ON THE PEOPLE WHO VISIT MUSEUMS.

WHEN EVALUATING A PROGRAM IN A MUSEUM CONTEXT, IT QUICKLY BECOMES APPARENT THAT TECHNIQUES EMPLOYED IN SUCH A MILIEU MUST BE ABLE TO ACCOMMODATE AN OUTCOME AS PERSONAL AS A MUSEUM EXPERIENCE. THIS IMPLIES THAT THE IMPACT OF A PROGRAM IS NOT NECESSARILY QUANTIFIABLE AND IS NOT IDENTICAL FOR ALL PARTICIPANTS. CONSEQUENTLY, NATURALISTIC/RESPONSIVE APPROACHES OFTEN ARE USED.

RESEARCHERS TRAINED IN CLASSICAL TECHNIQUES OF EVALUATION AT TIMES HAVE DENIGRATED THE"SOFT" METHODS EMPLOYED BY NATURALISTIC EVALUATORS. THIS ATTITUDE FREQUENTLY RESULTS FROM A MISUNDERSTANDING REGARDING THE DEGREE OF DIFFICULTY ENCOUNTERED WHEN ATTEMPTING TO MEASURE INTANGIBLES SUCH AS INDIVIDUAL AFFECT. IN ADDITION, RESEARCHERS EMPLOYED IN THE INVESTIGATION OF THESE MILIEUS MUST FACE THE HARSH REALITY THAT PROBLEMS OF MEASUREMENT INHERENT IN SUCH DISTINCT CONTEXTS ARE OFTEN INCOMPREHENSIBLE, NOT TO MENTION IRRELEVANT, TO INDIVIDUALS SEEKING PROFESSIONAL ASSISTANCE. UNFORTUNATELY, THE SAME CANNOT BE SAID FOR FELLOW EVALUATORS.

THE BASIC PROBLEM, THEN, INVOLVES THE NECESSITY OF RECONCILING TWO DIFFERENT WORLDS. ON THE ONE HAND, PROFESSIONAL STANDARDS AND SOUND EVALUATION PRACTICES MUST BE MAINTAINED WHILE, ON THE OTHER HAND, CLIENTS' NEEDS MUST BE MET AS COMPLETELY AS POSSIBLE.

DIFFICULTIES WITH NON-CLASSICAL APPROACHES

Problem Formulation

After conducting an evaluation in a naturalistic/responsive mode, this researcher became aware of several concerns which should be addressed by anyone who is considering using similar techniques in future studies. First the issue of problem formulation should be considered.

Given the exploratory nature of evaluations undertaken using naturalistic approaches, pre-set hypotheses are generally not appropriate. This does not mean that advance preparation and thought are not necessary. In fact, due to the ambiguity of the "hunt" itself, a great deal of groundwork (and even homework) is necessary. The nature of the rpoblem being explored should be understood as thoroughly as possible, although it is equally important that the evaluator maintain a degree of objectivity concerning the arena of investigation.

Also, the relative laxity in the determination of specific evaluation questions when working in a naturalistic mode should not be construed as an"easy out" by researchers eager to get on with the happy task of manipulating some variables. Rather, the very real issue of dealing with ambiguity must be faced. An evaluator should feel confident about his or her ability to commence a search when the exact questions to be addressed are less than completely formulated.

Iterative Framework

This leads to a closely-related issue in non-classical evaluation approaches. To wit, iterative problem formulation is a necessary

CONSIDERATION. THIS MEANS THAT QUESTIONS TO BE ADDRESSED BY THE EVALUATION SHOULD INITIALLY BE DETERMINED, RELEVANT DATA SHOULD BE GATHERED AND REPORTED BACK TO THE EVALUATION AUDIENCE, AND THEN THE ORIGINAL SET OF QUESTIONS SHOULD BE REFINED AND EXPANDED IF NECESSARY. TO ILLUSTRATE, INITIAL RESULTS OBTAINED IN THE STUDY AT THE TOLEDO MUSEUM OF ART ON VISITOR BEHAVIORS BROUGHT THE CURATORIAL STAFF TO A LEVEL WHERE THEY COULD POSE QUESTIONS WHICH WERE ACTUALLY OF GREATER IMPORTANCE TO THEM THAN WERE THE QUESTIONS WHICH WERE ORIGINALLY CHOSEN BY THE STAFF FOR INVESTIGATION.

SUCH A DYNAMIC APPROACH TO THE ESTABLISHMENT OF RESEARCH PARAMETERS UNDOUBTEDLY ENABLES AN EVALUATOR TO RESPOND TO CLIENT NEEDS WITH ALACRITY AS WELL AS EFFICACY. HOWEVER, THE POSSIBLE SIDE EFFECT OF FRUSTRATION SHOULD NOT BE TREATED LIGHTLY - SOME EVALUATORS HAVE ACTUALLY BEEN KNOWN TO RESIST (NOT TO MENTION RESENT) THE OBLIGATION TO SWITCH RESEARCH HORSES IN MID-CHI SQUARE.

EVALUATION AUDIENCE PARTICIPATION

WHILE ON THE SUBJECTS OF FORMULATION AND ITERATION, IT IS RECOMMENDED THAT THE EVALUATION AUDIENCE BE INCLUDED IN THE PROBLEM DETERMINATION STAGES TO A LARGE EXTENT. THIS IS ESPECIALLY IMPORTANT IN ATYPICAL FIELDS OF INVESTIGATION. IN ADDITION TO THE POSSIBLE LACK OF FAMILIARITY ON THE PART OF THE RESEARCHER WITH THE ENVIRONMENT TO BE STUDIED, THE RELEVANT STAFFS OFTEN HAVE MISCONCEPTIONS CONCERNING BOTH THE NATURE OF EVALUATION AND THE NATURE OF THEIR CLIENTS. ANY OF THESE SITUATIONS CAN PRESENT UNNECESSARY COMPLICATIONS TO THE EVALUATION PROCESS.

FOR EXAMPLE, IN THE MUSEUM ANALYSIS UNDERTAKEN BY THIS EVALUATOR, INTRODUCTORY PROBLEM FORMULATION STAGES PROVIDED THE CURATORIAL STAFF WITH A MORE REALISTIC VIEW OF THE KINDS OF INFORMATION THE EVALUATION COULD PROVIDE. ALSO, PERSONNEL BECAME BETTER ACQUAINTED WITH THE NATURE OF VISITORS TO THEIR PARTICULAR MUSEUM, AND THUS WERE ABLE TO FOCUS THEIR ATTENTIONS ON ISSUES PERTINENT TO THEIR UNIQUE ASSORTMENT OF CLIENTS.

BUT - JUST WHO WAS IT WHO ORIGINALLY SAID "A LITTLE KNOWLEDGE IS A DANGEROUS THING"? A REAL-LIFE ILLUSTRATION HAS BEEN KNOWN TO REAR ITS UGLY HEAD WHEN NONTRADITIONAL CONSUMERS OF EVALUATION DATA ARE EXPOSED TO RUDIMENTARY TECHNIQUES OF INQUIRY. AUDIENCES RELATIVELY UNEDUCATED REGARDING THE MORE MUNDANE, ALBEIT GENUINE, CONCERNS OF PROFESSIONAL RESEARCHERS OFTEN FALL VICTIM TO GRANDIOSE VISIONS OF THE INVESTIGATIVE OUTCOMES. TEMPERING SUCH EXPECTATIONS FOR THE MASSIVE AMOUNTS OF DATA TO BE GENERATED BY ONE LITTLE STUDY WITH THE REALITIES OF LIFE AS A HARRIED EVALUATOR, OR EVEN AS A JADED GRADUATE STUDENT ANXIOUS TO COMPLETE A DISSERTATION, MUST NOT BE AVOIDED.

OF COURSE, IF ONE'S CLIENTS TO NOT APPEAR TO APPRECIATE OR UNDERSTAND ONE'S EFFORTS TO SPARE THEM THE DISAPPOINTMENT OF UNMET EXPECTATIONS, IT WOULD NOT BE DEEMED IMPRUDENT TO REACQUAINT SAID CLIENTS WITH THE PROBABLY NONEXPANSIVE NATURE OF THE REMUNERATION PROFERRED FOR THE CONDUCT OF THE EVALUATION (THIS TECHNIQUE IS ESPECIALLY HEARTRENDING IN THE CASE OF THE GRADUATE STUDENT, WHOSE FLEDGLING EXPERTISE IS FREQUENTLY GIVEN BOTH LITERAL AND FIGURATIVE "FREE" REIN).

. . .

Classical Compromises

Perhaps the most persistent criticism of naturalistic/responsive approaches cited by researchers promoting classical evaluation techniques involves the almost wanton abandon with which traditional concerns are treated. While skills necessary for contending with potential threats to the internal and expernal validity of a study must be possessed by any evaluator who purports to be a professional, the ability to respond to the equal likelihood of encountering a less-than-ideal research situation must also be nurtured. In other words - sure it would be great to control for the possible contamination effects of instrumentation bias, or to make decisions regarding museum visitors in general based on the findings at this one building. Unfortunately, textbook case studies seem to have few living relatives in the workaday world of evaluation.

When a researcher is concerned with individual affective responses to a treatment by any definition, generalities are tenuous at best. Further complications arise when there are no "right" or "wrong" replies to the research questions. If all answers are equally valid, all must be given credence and must be duly reported.

This implies that those marvelous statistical techniques such as F-tests and multiple regressions which many researchers have come to know and love, and which lend themselves so well to both table and jargon generation, are, alas, inappropriate.

Since qualification rather than quantification is often the goal of evaluations in non-traditional settings, slightly more complicated and more time-consuming techniques of data reporting and

COLLECTION MUST BE EMPLOYED. NO SIMPLE TEST FOR SIGNIFICANCE BETWEEN TWO MEANS WILL BE POSSIBLE. IN FACT, EVEN THE HUMBLE CHI-SQUARE BECOMES RISKY IN MANY CASES.

SO WHERE DOES THIS LEAVE THE DEDICATED BUT INTREPID EVALUATOR WHO WISHES TO SATISFY BOTH CLIENT NEEDS AND PROFESSIONAL MANDATES? IN GENERAL, THE TASK AHEAD WILL INVOLVE ACTUAL VERBAL AND OBSERVATIONAL CONTACT WITH NUMEROUS CLIENTS, SO THE COMFORT AND ANONYMITY OF THE GENERATION OF COMPUTER OUTPUT MUST BE FORSAKEN. ALSO, MODIFICATIONS IN REPORT WRITING MUST BE MADE. SINCE QUALITATIVE DATA DO NOT LEND THEMSELVES EASILY TO CONDENSATION INTO TABLES AND FIGURES, ANOTHER SECURITY BLANKET MUST BE DISCARDED. INSTEAD, THE SPOKEN WORD IS GIVEN NEW CONSIDERATION, AND RESEARCHERS MUST REVIEW THE USE OF QUOTATION MARKS IN TEXT.

FINALLY, EXPECTATIONS ABOUT ADDING SIGNIFICANTLY TO THE BODY OF EVALUATION KNOWLEDGE SHOULD ALSO BE LOWERED. THE ABILITY TO GENERALIZE BACK TO THE SUBJECTS OF THE INDIVIDUAL STUDY IS SOMETIMES ALL ONE CAN HOPE FOR IN NON-TRADITIONAL SETTINGS. WHILE ENTIRE POPULATIONS ARE WAITING OUT THERE TO BE LABELED AND DISCUSSED WITHOUT ACTUALLY BEING STUDIED, NATURALISTIC EVALUATIONS ARE NOT THE PROPER VEHICLES FOR DOING SO.

PRECAUTIONS FOR POTENTIAL PROBERS

HAVING EARNED A NATURALISTIC EVALUATION BADGE AS A RESULT OF A MUSEUM STUDY, THIS RESEARCHER BELIEVES IT IS NOT PRESUMPTUOUS TO OFFER A BIT OF ADVICE TO THOSE CONSIDERING THE CONDUCT OF SIMILAR NON-

. . .

TRADITIONAL INVESTIGATIONS.

1. BEFORE ALL ELSE, THINK TWICE ABOUT WHAT YOU ARE ABOUT TO DO. AFTER ALL, YOU NOW KNOW THE JOB IS DANGEROUS BEFORE YOU TAKE IT.

2. IF YOU ARE STILL DETERMINED, NEXT REALISTICALLY ASSESS HOW WELL YOU LIKE PEOPLE IN GENERAL. THESE APPROACHES WILL REQUIRE MASSIVE AMOUNTS OF HUMAN CONTACT, SO MAKE SURE YOU ARE UP TO IT.

3. PRIOR TO FINAL ACCEPTANCE OF THE TASK, CAREFULLY ASCERTAIN YOUR ABILITY TO WORK WITH THE EVALUATION AUDIENCE. IT IS LIKELY EACH PARTY WILL NEED TO DEPEND ON THE OTHER FOR INFORMATION AND GUIDANCE, SO MUTUAL EASE OF COMMUNICATION IS ESSENTIAL.

4. UNDER THE HEADING OF SELF-PRESERVATION COMES THE RECOMMENDATION TO RECRUIT TRAINED ASSISTANTS WHENEVER POSSIBLE. THE AMOUNT OF TIME THAT IS REQUIRED FOR INSTRUMENT FORMULATION AND DATA COLLECTION OFTEN BECOMES FORMIDABLE, ALTHOUGH THIS IS NOT THE IMPRESSION ONE IS LIKELY TO HAVE AT THE EVALUATION'S OUTSET. INCIDENTALLY, THIS STEP SHOULD BE COMPLETED PRIOR TO THE INITIATION OF RESEARCH ACTIVITIES. ONCE THE TRUE NATURE OF THE PROBLEM IS REVEALED, ASSISTANTS WILL BE NEXT TO IMPOSSIBLE TO LOCATE.

5. PERHAPS MOST IMPORTANTLY, ALWAYS KEEP THE EVALUATION AUDIENCE IN MIND. MORE THAN LIKELY THEY HAVE NO INTEREST IN THE ASSUMPTIONS FOR F-TESTS, AND THEY COULD NOT DISTINGUISH BETWEEN 1-TAILED AND

2-TAILED TESTS IF THE ISSUE WERE TO ARISE. HOWEVER, THE PEOPLE FOR WHOM THE EVALUATION IS BEING CONDUCTED UNDOUBTEDLY CARE A GREAT DEAL ABOUT THEIR CLIENTS. THEREFORE, A PRIMARY CONSIDERATION SHOULD BE THE PROVISION OF DATA THAT WILL BE OF GREATEST USE TO THE AUDIENCE.

IN A DISCUSSION OF THE RELATIVE MERITS OF NATURALISTIC/RESPONSIVE EVALUATION TECHNIQUES IN NON-TRADITIONAL SETTINGS, THE MAJOR CONSIDERATION SEEMS TO INVOLVE ONE'S ABILITY TO COLLECT MEANINGFUL DATA UNDER LESS THAN IDEAL (LET ALONE ROBUST) CONDITIONS. ONCE THE LIMITATIONS OF SUCH STUDIES ARE RECOGNIZED AND ACCEPTED, THE ADVANTAGES CAN BE HIGHLIGHTED. PERHAPS THE GREATEST PAY-OFF OF THESE EVALUATIONS IS THAT THE KINDS OF DATA GATHERED AND REPORTS SUBMITTED ARE OFTEN OF IMMEDIATE AND PRACTICAL USE TO THE EVALUATION AUDIENCE. FOR INSTANCE, AS A DIRECT RESULT OF THIS RESEARCHER'S STUDY FOR THE TOLEDO MUSEUM OF ART, SEVERAL GALLERIES WITHIN THE MUSEUM WERE REPOSITIONED, AND NON-PRODUCTIVE FORMS OF INFORMATION DISSEMINATION WERE ELIMINATED. IN ADDITION, FUTURE SPECIAL EXHIBITS WILL INCLUDE PROVISIONS FOR EVALUATION OF EFFECTIVENESS.

EVEN THOUGH R.A. FISHER MIGHT NOT IDENTIFY THE TECHNIQUES EMPLOYED IN THE EVALUATIONS OUTLINED IN THIS PAPER AS BEING ACCEPTABLE RESEARCH TOOLS IN THE STRICT CLASSICAL SENSE, SUCH APPROACHES ARE NONETHELESS VALUABLE. IT BEHOOVES ALL EVALUATORS TO BECOME FAMILIAR WITH NON-TRADITIONAL FORMS OF INVESTIGATION, AS THE PRESENCE OF SUCH TECHNIQUES IS LIKELY TO BECOME INCREASINGLY PERVASIVE. AS POTENTIAL CLIENTS

. . .

FOR EVALUATIONS BECOME DIVERSIFIED IN BACKGROUNDS AND INTERESTS, RESEARCHERS MUST BE WILLING AND ABLE TO ADAPT TO THE DYNAMICS OF EACH UNIQUE SITUATION AS IT EXISTS IN AN INTANGIBLE AND NON-QUANTIFIABLE WORLD.